EXHIBIT C

UCR Pilot - Plan for 2000

Simona Cohen on 06/02/2000 at 14:27
Category: ISO

Unified Customer Reporting (UCR) Pilot - Plan for 2000

Summary
Project Title: Unified Customer Reporting (UCR) Pilot
Description (1-2 lines):
Following the UCR proof-of-concept in 1999, we'll develop a pilot to integrate and correlate SRM and EPP data sources. The integration is done by defining mappings from the data sources to a unified schema DTD and creating an XML representation of the unified data. The data is then visualized using the CREDiT applet or something similar, that allows to see information and monitor problems at a glance while preserving the ability to drill down to detailed information.

Proposed Research Leader: Simona Cohen
Proposed Global Services Leader: Gary Quesenberry

Problem Description (a paragraph or so):
In today's business environment, many applications need to use data that is warehoused in diverse data sources and repositories. The data is expressed in different formats and languages, retrieved in different access methods, and through different delivery vehicles. Moreover, new data sources may be added, and existing ones removed or changed frequently.
This problem is more and more common in areas where applications and databases were developed gradually over many years to supply increase demand in organisations for Information Systems (IS). Thus, it is critical with organisations that are using IS technology for many years.

IGS would like to develop one such application that includes two diverse performance data sources:
SRM (Server Resource Management) -- measures daily resources of servers such as CPU utilization, percent of memory used, number of users logged in, etc.
EPP (End-to-end Probe Performance) -- measures the response time of probes such as round-trip e-mail messages, access to specific web pages, etc.
The application will correlate and report on information from both of these sources. The report should allow examination of EPP data, along with details from SRM data, in order to analyze what caused poor performance of a probe. For example, probes work on servers, therefore reports would detail the performance of the servers that worked on the day the probe failed. Similarly, the report will investigate SRM data while using the EPP data to analyze how a probe performs at a specific load on a server.

The pilot will solve the problem of data integration for SRM and EPP, but will not preclude from adding other data sources in the future. Additionally, it will supply some generic components that may be used in data integration problems in other areas.

Global Services Business Opportunity (a paragraph or so):

Deliver quality services - Global Services and our customers require new system performance reports that include both server resources performance and probes performance. By looking at the different data sources and seeing how the information is related, they can get a better breakdown of the entire situation. The pilot will supply such reports in a novel way.

Cost reduction - The pilot will include some generic components in the back-end. This will allow faster integration of diverse data sources in the future, reduce development cost, and increase reuse and flexibility.

Global Services Business Impact (a paragraph or so):

Improved tools for customer reporting.
Improved market share in system management services.
Better competitiveness by having the ability to lower cost structure with better tools and more flexibility.

Proposed Research deliverables to Global Services (with impact of each):

An administrator application to map from data sources to the unified schema DTD.
impact: faster and easier creation of mappings by a visual application.
A back-end system including a Lookup Engine that integrates the diverse data sources and creates the unified data.
impact: generic infrastructure to ease data-integration problems.
A visualization applet that will report and analyze the unified data.
impact: leverage system management services.
Ongoing evaluation on report requirements and incorporation to our report applet.
impact: leverage the quality of our report.

A Few Key Proposed Milestones:

The proposed milestones depend on the following prerequisites:
Getting access to SRM and the new EPP database by March 15th.
Getting the revised schemas of SRM and EPP by March 15th.
Having real data for the pilot selected customer by April 30th. The data should include EPP probes that operate on servers which are monitored by SRM.

Here are the proposed milestones:
Q1
DOU document.
Requirements document for all 3 components: the administrator, the back-end system and the CREDiT visualization applet.
Technology transfer of the CREDiT visualization applet from Watson to Haifa.
Study the revised schemas of EPP and SRM. This depends on

prerequisite 2.
Establish connection to EPP and SRM data sources. This depends on prerequisite 1.

Q2
Design document for all 3 components: the administrator, the back-end system and the CREDiT visualization applet.
Development of all 3 components: the administrator, the back-end system and the CREDiT visualization applet.
Development of conversion functions for the data sources.
Study the pilot final schemas of EPP and SRM.
Decide on the pilot structure and select a customer.
Deliver an initial alpha version of the system with limited functionality. This strongly depends on all three prerequisites.

Q3
Development of all 3 components: the administrator, the back-end system and the CREDiT visualization applet.
Development of conversion functions for the data sources.
Deliver static reports with real data for the selected customer. This strongly depends on all three prerequisites.

Q4
Enhancements to all 3 components according to the customers feedback.
Adding customer isolation.
Adding more customers to the pilot.
Work on performance improvements to the back-end system.
Add dynamic (on-line) reports (optional).
Adding documentation.

Future development owner (who will continue with this project once research is finished):
IGS Outsourcing

Cost of project (per quarter and total):
1Q2000 - 9 PM
2Q2000 - 9 PM
3Q2000 - 9 PM
4Q2000 - 9 PM
Total - 36 PM

ROI (Per year after first deliverable for 5 years):

Details
Introduction (describe the problem, technically, as well as from a business perspective including how large the problem is):
During the first quarter of 1999, IGS has posed to IBM research the following problem:
"IGS is obligated to provide different kinds of data/information to different commercial accounts. Examples include measurement data, reports, billing information, trouble ticket data, inventory data, usage data, capacity data, etc. Today each piece of data is provided in different ways to different accounts. The data comes from different sources, is presented in different report formats, using different access methods, through different delivery vehicles, etc. And beyond

that, IGS itself has requirements to access similar data on its accounts for internal purposes."

Unified Customer Reporting (UCR) aims to solve the problem by providing a generic infrastructure for data integration of multiple heterogeneous correlated data sources. UCR is open, extensible, modifiable and web-based for maximum client access. It is based on multi-platform standards and IBM strategic infrastructure including WebSphere, XML technology, Servlets, Java. UCR presents the data as a structured text file in XML, so that it can be viewed and used by applications other than UCR applications. XML is rapidly becoming accepted as the standard for data interchange-both across the Web and between applications.

The Server Resource Management (SRM) data source and the End-to-end Probe Performance (EPP) data source are two separate data sources for system performance measurement. SRM data source keeps historical information on servers utilization for thousands of distributed servers on all of the major server platforms: Windows NT,Novell Netware, AIX, HP-UX, and Sun Solaris. In addition, SAP and Lotus Notes application trending are available for servers running those applications. EPP data source includes duration measurements of emulated end-user transactions. The transactions emulate application behaviour and assess the quality of application services from both a historical and real time perspective.

Each one of those data sources gives a different angle of the performance puzzle. By integrating and correlating the two data sources, we can give our customers a better breakdown of the entire situation. This integration will provide the ability to answer questions like what was the performance of the servers when the application (probe) failed? How changes to applications affect the servers? How applications perform at a specific load on a server? Which component in the server had a bottleneck when the server runs a specific application? Which component in the application had a bottleneck when it runs on some specific servers?

Following a demo in 1999, we are going to develop in 2000 a pilot to integrate the SRM and EPP data sources using a generic data integarion infrastructure. The unified data it then visualized using a novel data analyzing concept called LifeLines which is the concept behind the CREDiT applet.

Solution (describe what will be done and how it will solve the problem):

The solution includes a back-end and a front-end with XML data in the interface between them. The back-end includes a novel method and a middleware system for integrating diverse data sources using eXtensible Markup Language (XML). The front-end is based on LifeLines (sometimes called also TimeLines) that was developed by the Watson Research Lab and the Human Computer Interaction Lab at the University of Maryland. We'll describe each part separately.

The back-end consists of an Administrator application and a Lookup Engine. A unified schema represented in a Document Type Definition (DTD), and a repository of mappings are created to map from the data sources to the unified schema. Then, the Lookup Engine uses the repository of mappings to extract the relevant data from the data

sources and create the unified data. The unified data is represented in XML and complies to the unified schema DTD. All the complexity of accessing the data sources, retrieving the data, and correlating it is done by the Lookup Engine, and is transparent to the application which just uses the output of the Lookup Engine - the unified data.

In 1999 we defined a Unified Schema DTD for performance, called performanceML, and it was published externally in xml.org.

The following figure sketches the proposed system:

(Embedded image moved to file: pic12143.jpg)

The front-end includes a Java visualization tool for rapid interpretation of temporal categorical data and is based on Plaisant & Shneiderman LifeLines. It allows you to see information at a glance while preserving the ability to drill down to see detailed backup information. It presents the data in layers and allows you to discover patterns. When there is a need for viewing and categorizing a lot of data, traditinal human-interaction techniques may fall short. Long lists to scroll, clumsy searches, endless menus and lengthy dialogs will lead to user rejection. LifeLines include techniques to summarize, filter and present large amounts of information, leading us to believe that rapid access to needed data is possible with careful design.

The front-end includes the following features:

Each entity i.e. probe or server has its own shape (screen representation). We correlate one to the other by overlaying the servers shape over the probes shape.

The use of a colormap to highlight certain features of a presentation is a common technique in scientific visualization and image processing. Most often colormaps are used to apply color based on the value of a continuous variable, however the use of categorical colormaps has been studied as well and is used here. It allows to direct attention immediately to the items requiring immediate attention.

The fundamental concept is that the position and size of the horizontal bars (probes) provide a fast way to see what happened and the temporal relationships among things which have happened. Positioning on the Y axis is accomplished principally through categorization and sorting.

It includes a thumbnail overview and scrolling of the larger central visual representation. The thumbnail serves to direct the user's attention to appropriate parts of the display. The small rectangular box inside the thumbnail sketch performs the function of scroll bars.

It allows to drill-down to detailed information.

Deliverables (list the deliverables and describe for each what the value and impact will be...what will Global Services do with each deliverable):

An administrator application to map from data sources to the unified schema DTD.

impact: faster and easier creation of mappings by a visual application.

A back-end system including a Lookup Engine that integrates the diverse data sources and creates the unified data.

impact: generic infrastructure to ease data-integration problems.

A visualization applet that will report and analyze the unified data.

impact: leverage system management services.

Ongoing evaluation on report requirements and incorporation to our report applet.

impact: leverage the quality of our report.

Technical (what will be done technically and why is research doing it...new technical ground or is it proving something in the marketplace):
The technology proposed for the back-end is new in the marketplace. A related technology is Virtual DB from Enterworks, but this proposal goes beyond Virtual DB by taking advantage of the new emerging XML technology. The technology proposed for the front-end was used before in the healthcare domain, and is new to the system performance domain. This will need new technical ground.

The technical work to be executed by HRL:
Design/implement a Look-Up Engine for the back-end.
Design/implement the Administrator .
Implement conversion functions for the SRM and EPP data sources.
Design/implement the visualization applet.

The technical work to be executed by IGS:
Assistance in establishing connection to SRM and EPP databases.
Assistance in understanding the new SRM and EPP schemas.
Assistance in defining requirements/specifications.
Assistance in design review (optional).
Selecting the pilot environment and customer.
Assistance in deploying the pilot in Raleigh.

Owner (who will take long term responsibility for this work, and what is their involvement now):
IGS Outsourcing

Competitive Analysis (who else, inside as well as outside IBM has solved or is solving this problem and what are the strengths and weaknesses of this solution):
Traditional solutions to integrating multiple data sources are of two types. One type is to perform the data integration at the application level (2-tier solution). This solution couples the business logic with the data and makes the application development expensive and difficult to support, customize or change. The other type of data integration is to create a new data warehouse and copy the data from the original diverse data sources to the warehouse. This solution is heavy and not flexible to dynamic changes in the data sources. IBM DataJoiner is a product that employs this solution.

Development Plan...
Tasks (list all of the project tasks as well as the Research and Global Services person months for each):
Staffing (Who are the people performing the tasks, and what is their availability - 50%, 100% etc.)

Simona Cohen - 100%
Leonid Dubinsky - 60%
TBD - 100%
Hilit Grosberg - HC

Detailed Milestones (List all of the milestones for this project - at least one per month):

Assumption: Start date March 1st 2000

The proposed milestones depend on the following prerequisites:
Getting access to SRM and the new EPP database by March 15th.
Getting the revised schemas of SRM and EPP by March 15th.
Having real data for the pilot selected customer by April 30th. The data should include EPP probes that operate on servers which are monitored by SRM.

The prerequisites were reviewed and approved by Chris Molloy and Gary Quesenberry.
They are attainable based on the following additional assumptions posed by Gary:

Andy Frenkiel will be working with Chris Molloy to provide the latest EPP schema for incorporation on our MVS test system; so we expect that your access to the SRM and EPP database and revised schema, by March 15th, will be no problem
Your EPP contact for the project will be Chris Molloy (with dependency on Andy Frenkiel)
In addition, we plan to permit your access to the real data by April 30th; with the target scope to include the Watson Research Notes servers and data
We will use these milestones, as well, to gauge our own progress

The proposed milestones:

03/2000: DOU document.
03/2000: Requirements document for all 3 components: the administrator, the back-end system and the CREDiT visualization applet.
03/2000: Technology transfer of the CREDiT visualization applet from Watson to Haifa.
03/2000: Study the revised schemas of EPP and SRM. This depends on prerequisite 2.
03/2000: Establish a UCR server in Haifa with connection to EPP and SRM data sources. This depends on prerequisite 1.
04/2000: IGS review of the DOU document.
04/2000: IGS review of the requirements document.
04/2000: Design document for all 3 components: the administrator, the back-end system and the CREDiT visualization applet.
04/2000: Enhancing the CREDiT code.
05/2000: IGS review of the design document (optional).
05/2000: Study the pilot final schemas of EPP and SRM.
05/2000: Decide on the pilot structure and select a customer.
06/2000: Visualizing initial real data with CREDiT.
06/2000: Initial alpha version of the back-end, administrator and conversion functions. This strongly depends on all three prerequisites.
07/2000: IGS feedback for the alpha version system.
08/2000: Adding a new view to CREDiT.
08/2000: Testing the system with real data in Haifa.
09/2000: Deploying the system that includes static reports with real data for the selected customer to IGS.

This strongly depends on all three prerequisites.
10/2000: IGS feedback for the Q3 delivery.
10/2000: Adding customer isolation.
11/2000: Adding more customers to the pilot.
11/2000: Work on performance improvements to the back-end system.
12/2000: Add dynamic (on-line) reports (optional).
12/2000: Adding documentation.
12/2000: Deploying the system to IGS.

Support, Customization and Replication (How will this work continue to be executed once the project is complete):
Both the back-end and the front-end can be applied to other areas.

Development Environment (HW/SW used and where):
Development will be Java based and will use:
Java development environment
Windows/NT machines (400Mhz and above)
DB2 for Windows NT and DB2 Connect
Apache Web Server
WebSphere Application Server

Financial Plan (Convert all PYs into $/quarter for Research, Global Services and total):

	RES	GS	TOT
1Q	$30K	$90K	$120K
2Q	$30K	$90K	$120K
3Q	$30K	$90K	$120K
4Q	$30K	$90K	$120K
TOT	$120K	$360K	$480K

ROI Overview:

Project Initiated	1st Deliverable	First Year	Second Year	Third Year	Fourth Year	Each Year Thereafter
Time minus what (duration between start and 1st deliverable)?	Time 0	ROI $s	ROI $s	ROI $s	ROI $s	ROI $s

ROI Details (Logic behind estimate like -- Helpdesk agents receive repetitive calls on common problems that could easily be resolved by the customer if they had the correct procedure to provide educational instructions. These calls account for approximately 25-33% of the total volume received by the help desk. Putting this into perspective by applying it to South SDC call volumes (i.e. 2M per year), this correlates to about 500K-660K calls per year. Assuming a minimal 25% success factor on these calls, a rough resource savings would be 25 FTE(@ 500 calls per FTE/month) for the help desk):

Presentations (suitable for use with senior management to present all of the information described above - revise the status charts, as well as the status section above, on a monthly basis):

(See attached file: UCRPilot.PRZ)